New Cover Time Bounds for the Coalescing-Branching Random Walk on Graphs

Colin Cooper

King's College London, London, U.K.
colin.cooper@kcl.ac.uk

Tomasz Radzik

King's College London, London, U.K.
tomasz.radzik@kcl.ac.uk

Nicolas Rivera

University of Cambridge, Cambridge, U.K.
nicolas.rivera@cl.cam.ac.uk

August 15, 2019

Abstract

We present new bounds on the cover time of the *coalescing-branching random walk* process COBRA. The COBRA process, introduced in Dutta et al [9], can be viewed as spreading a single item of information throughout an undirected graph in synchronised rounds. In each round, each vertex which has received the information in the previous round (possibly simultaneously from more than one neighbour and possibly not for the first time), 'pushes' the information to k randomly selected neighbours. The COBRA process is typically studied for integer branching rates $k \geq 2$ (with the case $k = 1$ corresponding to a random walk). The aim of the process is to propagate the information quickly, but with a limited number of transmissions per vertex per round.

The COBRA cover time is the expected number of rounds until all vertices have received the information at least once. Our main results are bounds of $\mathcal{O}(m + (d_{\max})^2 \log n)$ and $\mathcal{O}(m \log n)$ on the COBRA cover time for arbitrary connected graphs with n vertices, m edges and maximum graph degree $d_{\max}$, and bounds of $\mathcal{O}((r^2 + r/(1-\lambda)) \log n)$ and $\mathcal{O}((1/(1-\lambda)^2) \log n)$ for r-regular connected graphs with the second largest eigenvalue λ in absolute value. Our bounds for general graphs are always $O(n^2 \log n)$, decreasing to $O(n)$ for constant degree graphs, while the best previous bound is $\mathcal{O}(n^{2.75} \log n)$. Our first bound for regular graphs applied to the lazy variant of the COBRA process is $O((r^2 + r/\phi^2) \log n)$, where ϕ is the conductance of the graph. The best previous results for the COBRA process imply for this case only a bound of $\mathcal{O}((r^4/\phi^2) \log^2 n)$.

To derive our bounds, we develop the following new approach to analysing the performance of the COBRA process. We introduce a type of infection process, which we call the *Biased Infection with Persistent Source* (BIPS) process, show that BIPS can be viewed as dual to COBRA and obtain bounds for COBRA by analysing the convergence of BIPS.

CCS Concepts: • **Mathematics of computing** ⟶ *Probabilistic algorithms*; • **Computing methodologies** ⟶ *Parallel algorithms*; *Distributed algorithms*.

Additional Key Words and Phrases: Random processes on graphs; epidemic processes; cover time; COBRA process

1 Introduction

Dutta et al. [9, 10] studied the following *coalescing-branching* random walk process for propagating information on a connected graph with n vertices and m edges. At the start of a round each vertex u which has just (in the previous round) received information 'pushes' this information to k randomly selected neighbours. Vertex u then stops sending the information until the time when it receives it again. At the end of a round, if a vertex receives information from two or more neighbours, then all these messages coalesce into one (we are assuming that all these messages carry the same information). The continuously repeated operations of COalescing messages and BRAnching by sending copies of the message to selected neighbours give the name COBRA to this process.

The COBRA process can be modelled as a particle process in the following way. At the start of each round, each existing particle divides into k particles (the branching factor). These particles then move independently to random neighbours. At the end of each round any particles which meet at a vertex coalesce to form a single particle.

The aim of the COBRA process is to rapidly propagate information to all vertices while limiting the number of transmissions per vertex per step, and without requiring that vertices store information for longer than one round. In the special case that $k = 1$, the COBRA process is a simple random walk, which achieves a low transmission rate but does not satisfy the fast propagation condition.

The main quantity of interest in information propagation processes (which is also studied in this paper) is the time taken to inform, or visit, all vertices. By analogy with a random walk, this quantity is referred to as the *cover time*. We next summarize the main known bounds on the cover time of the COBRA process and state our new bounds.

Dutta et al. [10] showed the following bounds for the case $k = 2$, which hold *with high probability* (w.h.p), that is, with probability at least $1 - n^{-c}$, for some constant $c > 0$, where n is the number of vertices in the graph. It is assumed that initially only one arbitrary (starting) vertex is informed.

(*i*) For the complete n-vertex graph, the cover time is $\Theta(\log n)$.

(*ii*) For regular constant-degree n-vertex expanders, the cover time is $\mathcal{O}(\log^2 n)$.

(*iii*) For the D-dimensional grid on n vertices, the cover time is $\mathcal{O}(n^{1/D}\text{polylog}\, n)$.

(*iv*) For n-vertex trees, the cover time time is $\mathcal{O}(n \log n)$, which is tight for the star graph.

Subsequently Mitzenmacher et al. [14] generalised bound (*ii*), improved bound (*iii*) and showed the first general bound for arbitrary connected graphs which was lower than the obvious $O(n^3)$ bound inherited from the (single) random walk case. More precisely, [14] shows the following bounds.

(*a*) For an r-regular n-vertex connected graph $G = (V, E)$, the cover time is $\mathcal{O}((r^4/\phi^2)\log^2 n)$, where ϕ is the *conductance* of G defined by

$$\phi = \min_{\emptyset \neq S \subsetneq V} \left\{ \frac{E(S : S^c)}{r \min\{|S|, |S^c|\}} \right\} \leq 1,$$

and $E(S : S^c)$ denotes the number of edges with one endpoint in S and the other in $S^c = V \setminus S$.

(*b*) For D-dimensional grid on n vertices, the cover time is $\mathcal{O}(D^2 n^{1/D})$, and this result is tight for D constant.

(*c*) For general n-vertex graphs, the cover time is $\mathcal{O}(n^{2.75} \log n)$.

The main contributions of this paper include the following new bounds on the cover time of the COBRA process. For arbitrary connected graphs, we show bounds $\mathcal{O}(m + (d_{\max})^2 \log n)$ and $\mathcal{O}(m \log n)$, where m is the number of edges and $d_{\max}$ is the maximum degree of a vertex. Both these bounds are always $\mathcal{O}(n^2 \log n)$, but become as small as $O(n)$ for constant-degree graphs (the first bound), significantly improving the previous bound (*c*).

For r-regular connected graphs, we show bounds $\mathcal{O}\big((r^2 + r/(1-\lambda))\log n\big)$ and $\mathcal{O}((1/(1-\lambda)^2)\log n)$, where $\lambda = \lambda(G) = \max\{\lambda_2, |\lambda_n|\}$ and $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$ are the eigenvalues of the transition matrix

of the random walk. These two bounds coincide when $1-\lambda = \Theta(1/r)$. Since $1-\lambda_2 \geq \phi^2/2$ (see [15]), if $\lambda_2 \geq |\lambda_n|$, then our first bound is tighter than the $\mathcal{O}((r^4/\phi^2)\log^2 n)$ bound given in [14] by a factor of $\Omega(r^2 \log n)$. We can impose the condition $\lambda_2 \geq |\lambda_n|$ by considering the *lazy* COBRA process, which in the branching part of each round allows each particle to choose with probability 1/2 to stay in its current vertex (independently of the choices of other particles).

As an example, consider the hypercube with $n = 2^d$ vertices, which is a regular graph of degree $r = \log n$ with both the conductance ϕ and the value of $1-\lambda_2$ of the order $\Theta(1/\log n)$ (the hypercube is bipartite, so $\lambda_n = -1$). The results in [14] imply an $O(\log^8 n)$ bound on the COBRA cover-time for the hypercube. The same bound applies for the lazy COBRA process and no better bound was known prior to our work. Our results give an $O(\log^3 n)$ bound on the cover time for the hypercube by the lazy COBRA process.

Observe that $\max\{\log_2 n, \mathrm{Diam}(G)\}$[1] , where $\mathrm{Diam}(G)$ is the diameter of a graph G, is an obvious lower bound on the cover time of the COBRA process with branching factor $k = 2$. Indeed, the number of visited vertices at most doubles in each round and a vertex at distance i from the starting vertex cannot be visited earlier than in round i. Thus our bounds give the first asymptotically optimal bound $\Theta(\log n)$ for the COBRA cover time for expanders (regular graphs with $1-\lambda$ greater than a positive constant) and the first asymptotically optimal bound $\mathcal{O}(n)$ for constant degree graphs.

Our main tool in analysing the COBRA process is a duality relation between this process and a particular discrete epidemic process, which we call the *biased infection with persistent source* (BIPS) process. This process is interesting in its own right and establishing the duality between COBRA and BIPS should be consider as one of the main contributions of this paper. The BIPS process is described as follows. A fixed vertex v is the source of an infection and remains permanently infected. At each round, each vertex u other than v, randomly selects k neighbours. In the next round vertex u will be considered as infected if, and only if, at least one of the neighbours selected in the current round is infected. Because of the persistent source, there will be eventually a round when all vertices will be infected, and they will remain infected in all subsequent rounds. The question is how quickly this will happen. It turns out that it is easier to work with the BIPS process than with the COBRA process and bounds on the *infection time* of BIPS can be directly translated into bounds on the cover time of COBRA.

The preliminary versions of the results included in this paper were presented in [6] and [7]. In [6] we introduced the BIPS process and derived a weaker $\mathcal{O}((1/(1-\lambda)^3)\log n)$ bound on the COBRA cover time for regular graphs. In [7] we introduced an additional aspect of our proof methodology: viewing BIPS as a sequential process, when in each round the vertices make their decision not simultaneously but one by one in some arbitrarily fixed order. This did not change the nature of BIPS in any way, as the sequential view of BIPS was only a proof technique, but allowed us to derive the bounds $\mathcal{O}(m + (d_{\max})^2 \log n)$ and $\mathcal{O}\big((r^2 + r/(1-\lambda))\log n\big)$ on the COBRA cover time. In this paper we put together all these results, include detailed proofs and derive two new COBRA cover-time bounds: $\mathcal{O}(m \log n)$ for general graphs and $\mathcal{O}((1/(1-\lambda)^2)\log n)$ for regular graphs.

Very recently and subsequently to our work, Berenbrink et al. [2] improved the bounds for the (original, not lazy) COBRA process on r-regular graphs as follows:

(*a*) For an r-regular n-vertex connected graph $G = (V, E)$, the cover time is $\mathcal{O}((1/\phi)\log n)$, where ϕ is the conductance of G.

(*b*) For an r-regular n-vertex connected graph $G = (V, E)$, the cover time is $\mathcal{O}(n \log n)$.

These results were obtained in [2] by refining the analysis of the BIPS process which we introduced in [6, 7] and detail in this paper.

Spreading information using the COBRA process has some similarities with broadcasting in the randomized push-based gossiping model [4, 12, 13], where in each round each informed vertex sends the message to a random neighbour. In the gossiping model each vertex has to remember whether it is informed and keeps sending out messages in each step from the time it has been informed. In that respect COBRA may be considered a simpler process for information disseminaiton. The COBRA process, however, is considerably harder to analyze since the set of the vertices sending out messages is not monotonic with the increasing time step.

[1] In this paper $\log x$ stands for the natural logarithm.

2 Definitions and formal statements of our results

Consider a connected graph $G = (V, E)$ with n vertices and m edges, an integer $k \geq 1$, and a subset of vertices $W \subseteq V$. The COBRA process, with starting set W and branching factor k, is the set-process $(W_t)_{t\geq 0}$ with $W_0 = W$, and the set W_{t+1} generated as follows. Each vertex $v \in W_t$ independently chooses k neighbours uniformly at random with replacement. Denote such a set of neighbours by $Y(v)$, then, W_{t+1} is defined by $W_{t+1} = \bigcup_{v\in W_t} Y(v)$. Note that a vertex in W_t does not necessarily belong to W_{t+1} (sets W_t are not monotonic with the increasing time step t). We can think of W_t as a set of vertices carrying a piece of information, and then each of them passes the information to k random neighbours to generate the new set of informed vertices W_{t+1}. For $W_0 = \{u\}$, let $\mathbf{cov}(u) = \min\{T \geq 0 : \bigcup_{t=0}^{T} W_t = V\}$ be the first round such that each vertex has been informed at least once. We proceed to state formally the main results of this paper.

Theorem 2.1. *Let G be a connected graph with n vertices, m edges and maximum vertex degree $d_{\max}$. Then for the COBRA process with branching factor $k = 2$ and for each $u \in V$, w.h.p. and in expectation $\mathbf{cov}(u)$ is*

$$\mathcal{O}(m + (d_{\max})^2 \log n), \tag{1}$$

and

$$\mathcal{O}(m \log n). \tag{2}$$

For a connected r-regular graph G with adjacency matrix $A(G)$, let $P = A(G)/r$ denote the transition matrix of the random walk on G. Let $\lambda_1 = 1 > \lambda_2 \geq \ldots, \geq \lambda_n \geq -1$ be the eigenvalues of matrix P ordered in a non-increasing sequence. Let $\lambda = \lambda(G) = \max_{i=2,\ldots,n} |\lambda_i| = \max\{\lambda_2, |\lambda_n|\}$ be the second largest eigenvalue in absolute value. Our second result gives a bound on the cover time of COBRA for regular graphs in terms of the *eigenvalue gap* $1 - \lambda$, and the degree r.

Theorem 2.2. *Let G be a connected r-regular n-vertex graph with eigenvalue gap $1 - \lambda \geq c\sqrt{(\log n)/n}$ for a suitably large constant c. Then for the COBRA process with branching factor $k = 2$ and for each $u \in V$, w.h.p. and in expectation $\mathbf{cov}(u)$ is*

$$\mathcal{O}\left(\left(\frac{r}{1-\lambda} + r^2\right) \cdot \log n\right), \tag{3}$$

and

$$\mathcal{O}\left(\frac{\log n}{(1-\lambda)^2}\right). \tag{4}$$

We note that for a connected graph, $1 - \lambda > 0$ if and only if the graph is not bipartite. For regular bipartite graphs, and more generally for regular graphs which are either bipartite ($\lambda_n = -1$) or close to bipartite (meaning for us that $|\lambda_n| > \lambda_2$, so $1 - \lambda < 1 - \lambda_2$), we can consider the *lazy* COBRA process which allows each vertex to select itself with probability $1/2$. Rather than making this change of the transition probabilities, we can equivalently modify the graph by adding at each vertex r self-loops. The lazy COBRA process on the r-regular graph G is the same as the (original, non-lazy) COBRA process on this modified $(2r)$-regular graph $G^{(L)}$. The i-th largest eigenvalue of the transition matrix $P^{(L)}$ of the random walk on graph $G^{(L)}$ is equal to $\lambda_i^{(L)} = (\lambda_i + 1)/2$, where λ_i is the i-th largest eigenvalue of the transition matrix P of the random walk on graph G. Thus all eigenvalues of matrix $P^{(L)}$ are non-negative, $\lambda^{(L)} = \lambda_2^{(L)} = (\lambda_2 + 1)/2$, and $1 - \lambda^{(L)} = (1 - \lambda_2)/2$. By applying Theorem 2.2, which holds for regular graphs with multiple edges and self-loops, to graph $G^{(L)}$, we can conclude that the same bounds (3) and (4) but with λ replaced with λ_2 hold for the lazy COBRA process on graph G.

We remark that all our results, as well as the main previous results, are given for the case $k = 2$. Note, however, that for $k > 2$, the cover time can only be smaller (as the process informs more vertices per round), so any upper bound on the cover time for the case $k = 2$ is also an upper bound for any $k \geq 2$.

Theorems 2.1 and 2.2 state that the given bounds hold w.h.p. and in expectation. It suffices, however, to prove that they hold w.h.p. since the same asymptotic bounds will then hold also in expectation by the following simple argument. Consider restarting the COBRA process after each run of T steps from any vertex in the current set W_T, if not all vertices have been covered yet. Assuming that each run of T steps fails to cover the whole graph with probability at most $1/n^c$, for some constant $c > 0$, then the expected cover time is at most $\sum_{i\geq 0}(1/n^{ci})T = O(T)$.

The COBRA process can be seen as a type of multiple random walks process, so it is tempting to try to analyse it using techniques developed for such processes. Previous work on multiple random walks include [1, 3, 5, 11], where cover times were analysed for various classes of graphs. The analyses of the COBRA process given in Dutta et al. [10] and Mitzenmacher et al. [14] use a number of tools from multiple random walks, but the applicability of those tools turns out to be limited because the random walks in COBRA are highly dependent. In this regard, we propose an alternative approach. Instead of directly analysing the COBRA walks, we analyse a related epidemic process, called BIPS. We show that the BIPS and the COBRA process are dual under time reversal, and thus properties of one process can be obtained by studying related properties in the other process.

Biased Infection with Persistent Source (BIPS): Consider a connected graph $G = (V, E)$, an integer $k \geq 1$ and a special *source* vertex $v \in V$. We consider the set-process $(A_t)_{t\geq 0}$ defined as follows. Let $A_0 = \{v\}$. Given A_{t-1}, in round $t \geq 1$ of the process, each vertex $u \in V$ other than the source v selects independently and uniformly with replacement k neighbours, and becomes a member of A_t if and only if at least one of the k selected neighbours is in A_{t-1}. Additionally, $v \in A_t$ for all $t \geq 0$. We view vertex v as the source of infection and set A_t as the infected set (the set of infected vertices) at time t (that is, at the end of round t). Observe that by putting v in all sets A_t we assume that the source v remains always infected. Other vertices are infected at the end of round t, if and only of, they have contacted in this round at least one neighbour which was infected at the end of the previous round. Thus a vertex other than the source remains infected, if it keeps contacting infected vertices.

The BIPS process is a discrete epidemic process of the SIS (Susceptible-Infected-Susceptible) type, in which vertices (other than the source v) refresh their infected state at each step by contacting k randomly chosen neighbours. The presence of a persistent (or corrupted) source means that, almost surely, all vertices of the underlying graph are eventually infected in the same round and remain infected in all subsequent rounds. The BIPS process is of independent interest since in the context of epidemics, certain viruses exhibit the property that a particular host can remain persistently infected.

Our main results for the COBRA process follow from the duality relationship between COBRA and BIPS stated in the theorem below. When referring to the probabilities of various events, to indicate the process (COBRA or BIPS), we use the notation $\mathbf{P}(\cdot)$ for the probabilities in the BIPS process, and $\widehat{\mathbf{P}}(\cdot)$ in the COBRA process. For the COBRA process and any vertex $x \in V$, let $\mathrm{Hit}(x) = \min\{t \geq 0 : x \in W_t\}$ be the hitting time of x, that is, the step when vertex x receives information for the first time. The theorem relates the event of informing (hitting) a vertex v by the COBRA process started from a set W and the event of infecting at least one vertex in set W by the BIPS process with source v. More precisely, the theorem states that the probability that v is not informed by COBRA by time t is equal to the probability that no vertex in W is infected by BIPS at time t.

Theorem 2.3. *Let G be a connected graph. Consider a COBRA process $(W_t)_{t\geq 0}$ and a BIPS process $(A_t)_{t\geq 0}$, both with the same parameter $k \geq 1$. Then for each $t \geq 0$, for each $v \in V$ and $W \subseteq V$ we have*

$$\widehat{\boldsymbol{P}}(\mathrm{Hit}(v) > t \mid W_0 = W) \;=\; \boldsymbol{P}(W \cap A_t = \emptyset \mid A_0 = \{v\}). \tag{5}$$

The infection (or completion) time $\mathbf{infec}(v) = \min\{T \geq 0 : A_t = V\}$ of the BIP process with the source vertex v is the first time when all vertices are infected. For any two vertices $u, v \in V$, and any $t \geq 0$, applying Theorem 2.3 with $W_0 = \{u\}$ gives

$$\begin{aligned} \widehat{\mathbf{P}}(\mathrm{Hit}(v) > t \mid W_0 = u) &= \mathbf{P}(u \notin A_t \mid A_0 = \{v\}) \\ &\leq \mathbf{P}(A_t \neq V \mid A_0 = \{v\}) \\ &= \mathbf{P}(\mathbf{infec}(v) > t). \end{aligned} \tag{6}$$

Let p be an upper bound on the probability that the BIPS process does not finish within time t, that is, $\mathbf{P}(\mathbf{infec}(v) > t) \leq p$, for each $v \in V$. Then the union bound and (6) give the following bound on the probability that the cover time of the COBRA process started from a vertex $u \in V$ does not finish within time T.

$$\widehat{\mathbf{P}}(\mathbf{cov}(u) > t) \;\leq\; \sum_{v\in V} \widehat{\mathbf{P}}(\mathrm{Hit}(v) > t \mid W_0 = u) \;\leq\; \sum_{v\in V} \mathbf{P}(\mathbf{infec}(v) > t) \;\leq\; pn. \tag{7}$$

Using (7), we get the bounds on the cover time of the COBRA process stated in Theorems 2.1 and 2.2 from the bounds on the infection time of the BIPS process stated in the following two theorems. Note that while the time bounds in Theorems 2.4 and 2.5 are the same as in Theorems 2.1 and 2.2, there is a difference in the probability bounds. In Theorems 2.4 and 2.5 we have to give explicit bounds on the probabilities of completing the BIPS process within the stated time to have room for that additional factor of n in (7), which the union bound introduces when we move from BIPS to COBRA .

Theorem 2.4. *Let G be a connected graph with n vertices, m edges and maximum vertex degree $d_{\max}$. For every $v \in V$, the infection time* $\mathbf{infec}(v)$ *of the BIPS process with $k = 2$ is*

$$\mathcal{O}(m + (d_{\max})^2 \log n), \tag{8}$$

and

$$\mathcal{O}(m \log n), \tag{9}$$

with probability at least $1 - \mathcal{O}(1/n^3)$.

Theorem 2.5. *Let G be a connected r-regular n-vertex graph with $1 - \lambda \geq c\sqrt{\log n/n}$ for a suitably large constant c. Then for every $v \in V$, the infection time* $\mathbf{infec}(v)$ *of the BIPS process with $k = 2$, is*

$$\mathcal{O}\left(\left(\frac{r}{1-\lambda} + r^2\right)\log n\right) \tag{10}$$

and

$$\mathcal{O}\left(\frac{\log n}{(1-\lambda)^2}\right) \tag{11}$$

with probability at least $1 - \mathcal{O}(1/n^3)$.

Unlike in the case of spreading information in the gossiping model, both the sequence $(A_t)_{t\geq 0}$ of the BIPS process and the sequence $(W_t)_{t\geq 0}$ of the COBRA process are not monotonic with the increasing time step, that is, sets A_t and W_t are not necessarily subsets of A_{t+1} and W_{t+1}. What makes the BIPS process easier to analyse than the COBRA process is the clear notion of the completion time, which does not refer to the history of the process: compare the definitions of $\mathbf{cov}(v)$ and $\mathbf{infec}(v)$. The BIPS process is completed at the first step T when $A_T = V$. Then for all $t \geq T$, $A_t = V$. To know that the COBRA process has completed by step t, we have to know that $W_0 \cup W_1 \cup \cdots W_t = V$.

The rest of this paper is as follows. First, in Section 3, we prove the duality relationship between COBRA and BIPS stated in Theorem 2.3. Then, in Section 4, we prove Theorem 2.4. To prove part (8) of this theorem, we view BIPS as a sequential process as mentioned in Introduction. Finally, in Section 5, we analyse the BIPS process on regular graphs and prove Theorem 2.5. This analysis is done in two steps. We first analyse infections with a large number of infected vertices, and then infections with a small number of infected vertices.

Additional notation. For $u \in V$, we denote by $N(u)$ the neighbourhood of u and, in general, for $A \subseteq V$, we define $N(A) = \bigcup_{u\in A} N(u)$. For $u \in V$, we define the degree $d(u)$ of u as the number of edges adjacent to u, counting the multiplicities of multiple edges and self-loops, if present. We denote $d_{\max} = \max\{d(u) : u \in V\}$. Given $X \subseteq V$, the degree $d(X)$ of set X is defined as the sum of the degrees of the vertices in X, and $d_X(u)$ is defined as the number of edges adjacent to vertex u which have the other end in set X.

3 Duality Between COBRA and BIPS Processes: Proof of Theorem 2.3

Recall that $(W_t)_{t\geq 0}$ and $(A_t)_{t\geq 0}$ denote the COBRA and BIPS processes, respectively, and that we use the notation $\mathbf{P}(\cdot)$ for probabilities in the BIPS process, and $\widehat{\mathbf{P}}(\cdot)$ for probabilities in the COBRA process. Moreover, to simplify notation, we may write "$A_0 = v$" for the frequently appearing condition "$A_0 = \{v\}$," which says that v is the source vertex of BIPS.

Observe that the claim of Theorem 2.3 is trivial if $v \in W$, since the probabilities on both sides in (5) are equal to 0. We assume therefore that $v \notin W$ and proceed by induction on t. For $t = 0$, the claim is true because both probabilities are equal to 1. Assuming that the claim is true for a fixed $t \geq 0$, we prove that it must be also true for $t + 1$. Thus for the induction step we consider the first $t + 1$ steps of the COBRA process and the first $t + 1$ steps of the BIPS process.

The idea behind the proof is to establish a correspondence between the *last*, $(t + 1)$-st step of BIPS and the *first* step of COBRA. At these two steps the vertices perform the same basic actions: each vertex selects k neighbours randomly, uniformly and independently of the choices of other vertices. One can think of the correspondence between these two steps as the vertices using the same choices of neighbours in both steps (coupling). Our formal proof, however, does not need this assumption, relying instead simply on the fact that the probabilities of selecting the same sets of neighbours are the same in both these steps. Unrolling the induction further, we make steps $t, t - 1, \ldots, 1$ in BIPS correspond to steps $2, 3, \ldots, t + 1$ in COBRA, respectively. We proceed to the formal proof.

Consider the BIPS process at step $t + 1$. Denoting by $X(u)$ the random multi-set of neighbours of size k chosen by a vertex $u \in V \setminus \{v\}$, we have $u \in A_{t+1}$, if and only if, $A_t \cap X(u) \neq \emptyset$. For convenience, we set $X(v) = \{v\}$. For a subset $S \subseteq V$, define $X(S) = \bigcup_{u \in S} X(u)$. The following event equalities hold.

$$\{S \cap A_{t+1} = \emptyset\} \;=\; \{\forall u \in S : X(u) \cap A_t = \emptyset\} \quad = \quad \left\{ \left(\bigcup_{u \in S} X(u) \right) \cap A_t = \emptyset \right\} \;=\; \{X(S) \cap A_t = \emptyset\}.$$

Therefore,

$$\begin{aligned} \mathbf{P}(W \cap A_{t+1} = \emptyset \mid A_0 = v) &= \mathbf{P}(X(W) \cap A_t = \emptyset \mid A_0 = v) \\ &= \sum_{B \subseteq V} \mathbf{P}(B \cap A_t = \emptyset, X(W) = B \mid A_0 = v) \\ &= \sum_{B \subseteq V} \mathbf{P}(B \cap A_t = \emptyset \mid A_0 = v)\, \mathbf{P}(X(W) = B). \end{aligned} \tag{12}$$

The last equality above follows from the fact that the assumptions of the model imply that at step $t + 1$, for any fixed sets $S \subseteq V$ and $B \subseteq V$, the event $\{X(S) = B\}$ (that is, the event that the set of the vertices selected at step $t + 1$ by the vertices in the set S is the set B) is independent of the random sets A_τ, $\tau = 0, 1, \ldots, t$, and thus also independent of the event $\{B \cap A_t = \emptyset\}$.

Consider now step 1 of the COBRA process, and recall that $W_0 = W$. For any vertex $u \in W$, let $Y(u)$ be the random multi-set of neighbours of u of size k chosen by u in this step of the COBRA process, and define $Y(W) = \bigcup_{u \in W} Y(u)$. Observe that $W_1 = Y(W)$. As $v \notin W$, then for any $u \in W$ and any $B \subseteq V$,

$$\widehat{\mathbf{P}}(Y(u) = B) \;=\; \mathbf{P}(X(u) = B). \tag{13}$$

The variables $X(u)$, for $u \in W$, are independent (since vertices choose their neighbours independently) and the same holds for the variables $Y(u)$. Thus, for any $B \subseteq V$,

$$\mathbf{P}(X(W) = B) = \widehat{\mathbf{P}}(Y(W) = B). \tag{14}$$

Finally, we continue (12) in the following way.

$$\begin{aligned} \mathbf{P}(W \cap A_{t+1} = \emptyset \mid A_0 = v) &= \sum_{B \subseteq V} \mathbf{P}(B \cap A_t = \emptyset \mid A_0 = v)\, \mathbf{P}(X(W) = B) \\ &= \sum_{B \subseteq V} \widehat{\mathbf{P}}(\mathrm{Hit}(v) > t \mid W_0 = B)\, \mathbf{P}(X(W) = B) \\ &= \sum_{B \subseteq V} \widehat{\mathbf{P}}(\mathrm{Hit}(v) > t \mid W_0 = B)\, \widehat{\mathbf{P}}(Y(W) = B) \\ &= \sum_{B \subseteq V} \widehat{\mathbf{P}}(\mathrm{Hit}(v) > t + 1 \mid W_1 = B, W_0 = W)\, \widehat{\mathbf{P}}(Y(W) = B) \\ &= \widehat{\mathbf{P}}(\mathrm{Hit}(v) > t + 1 | W_0 = W). \end{aligned}$$

The second equality above follows from the induction hypothesis. The third equality follows from (14). Thus the claim in the theorem is true for $t+1$, so by induction it is true for any $t \geq 0$.

4 Analysis of the BIPS process

4.1 Proof of Bound (9) in Theorem 2.4

In this section, we consider the BIPS process with source v on any connected graph, not necessarily regular and possibly with multiple edges and self-loops. To study the BIPS process, instead of tracking the infected set $(A_t)_{t\geq 0}$, we will track its degree, $d(A_t)$. For a given infected set $A \subseteq V$, we define the subsets of vertices $B = B(A)$ and $C = C(A)$ as

$$\begin{aligned} B &= \{u \in V : N(u) \subseteq A\}, \\ C &= (N(A) \cup \{v\}) \setminus B. \end{aligned} \tag{15}$$

The set B contains all vertices, possibly with the exception of the source vertex v, that will be surely infected in the next round, because all their neighbours are infected. On the other hand, and with the exception of v, the set C contains those vertices that may be infected by chance in the next round, because they have at least one infected neighbour (but not all neighbours are infected). The source vertex v is either in set B or in set C. Notice that if a vertex does not belong to $B \cup C$, then it is surely not infected in the next round.

For an infected set A_t, we denote by B_t and C_t the associated sets $B(A_t)$ and $C(A_t)$, respectively.

Lemma 4.1. *The event $\{A_t \neq V\}$ implies the event $\{C_t \neq \emptyset\}$.*

Proof. Suppose the source v does not belong to B_t, so $v \in C_t$ and C_t is not empty. If $v \in B_t$ then $N(v) \subseteq A_t$. Consider $u \in A_t^c \equiv V \setminus A_t$ and a path P joining v with u. Suppose $P = (v_1, \ldots, v_l)$ with $v_1 = v$, and $v_l = u$. Let j be the minimum index such that $v_j \in A_t^c$. Since $v \cup N(v) \subseteq A_t$, then $j \geq 3$. From the definition of j, it holds that $v_{j-2} \in A_t$ and $v_{j-1} \in A_t$ but $v_j \in A_t^c$. Then v_{j-1} has at least one neighbour in A_t and at least one neighbour in A_t^c, so $v_{j-1} \in C_t$ and C_t is not empty. □

We analyse the difference $d(A_{t+1}) - d(A_t)$ using sets B_t and C_t. Observe the following relation which refers to sets A_t, B_t, and C_t.

$$d(B_t) + \sum_{u \in C_t} d_{A_t}(u) = \sum_{u \in V} d_{A_t}(u) = \sum_{u \in A_t} d(u) = d(A_t). \tag{16}$$

The first equality holds because only vertices in B_t and C_t have neighbours in A_t, and also because all neighbours of vertices in B_t belong to A_t.

Let $X_{\tau,u}$ be the indicator variable taking value 1 if and only if $u \in A_\tau$. We have $X_{t+1,u} = 1$ for all $u \in B_t$, and $X_{t+1,u} = 0$ for all $u \in V \setminus (B_t \cup C_t)$. Therefore,

$$\begin{aligned} d(A_{t+1}) &= \sum_{u \in V} d(u) X_{t+1,u} \\ &= \sum_{u \in B_t} d(u) + \sum_{u \in C_t} d(u) X_{t+1,u} \\ &= d(B_t) + \sum_{u \in C_t} d(u) X_{t+1,u} \\ &= d(A_t) - \sum_{u \in C_t} d_{A_t}(u) + \sum_{u \in C_t} d(u) X_{t+1,u} \qquad (17) \\ &= d(A_t) + \sum_{u \in C_t} (d(u) X_{t+1,u} - d_{A_t}(u)), \qquad (18) \end{aligned}$$

where Equation (17) follows from Equation (16). To simplify notation, we will use $d_A(u)$ for $d_{A_t}(u)$.

Lemma 4.2. *Let $A \subseteq V$ and $v \in A$. Then*

$$\mathbf{E}(d(A_{t+1}) - d(A_t) \mid A_t = A) = \sum_{u \in C_t} \left(d_A(u) - \frac{d_A(u)^2}{d(u)} \right) \geq \frac{|C_t|}{2}.$$

Proof. If $A = V$, then both sides above are equal to 0. Assume therefore that $A \neq V$. Given $A_t = A$, we have that the random variables $X_{t+1,u}$, for $u \in V$, are independent of each other. Moreover, for $u \neq v$, it holds

$$\mathbf{P}(X_{t+1,u} = 1 | A_t = A) = 1 - \left(1 - \frac{d_A(u)}{d(u)}\right)^2, \tag{19}$$

because $X_{t+1,u} = 1$, if and only if u does not choose both neighbours outside A_t. Thus for $u \in C_t \setminus \{v\}$, using (19),

$$\begin{aligned} \mathbf{E}(d(u)X_{t+1,u} - d_A(u) | A_t = A) &= d(u)\left(1 - \left(1 - \frac{d_A(u)}{d(u)}\right)^2\right) - d_A(u) \\ &= d(u)\left(2\frac{d_A(u)}{d(u)} - \left(\frac{d_A(u)}{d(u)}\right)^2\right) - d_A(u) \\ &= d_A(u)\left(1 - \frac{d_A(u)}{d(u)}\right) \geq 1 - \frac{1}{d(u)} \geq \frac{1}{2}. \end{aligned} \tag{20}$$

In the last two inequalities, we use that for $u \in C_t \setminus \{v\}$ it holds $0 < d_A(u) < d(u)$, thus $d(u) \geq 2$, and that $d_A(u)(1 - d_A(u)/d(u))$ is minimised when $d_A(u)$ is equal to 1 or $d(u) - 1$. If $v \in C_t$, then $d(v)X_{t+1,v} - d_{A_t}(v) = d(v) - d_{A_t}(v) \geq 1$, because $d(v) > d_A(v)$, or otherwise v would belong to B_t and not to C_t. Thus for all $u \in C_t$,

$$\mathbf{E}(d(u)X_{t+1,u} - d_A(u) | A_t = A) \geq \frac{1}{2}. \tag{21}$$

Summing up (21) over all $u \in C_t$, and by Equation (18), the following holds.

$$\mathbf{E}(d(A_{t+1}) - d(A_t) \mid A_t = A) = \mathbf{E}\left(\sum_{u \in C_t} (d(u)X_{t+1,u} - d_{A_t}(u)) \mid A_t = A \right) \geq \frac{|C_t|}{2}.$$

□

Recall that $\mathbf{infec}(v)$ denotes the time it takes the BIPS process with source v to infect the whole graph G.

Theorem 4.3. *Let G be a connected graph with n vertices and m edges. Then for every $v \in V$, it holds that $\mathbf{E}(\mathbf{infec}(v)) = \mathcal{O}(m)$.*

Proof. Consider any $T \geq 1$, then by Lemma 4.2, it holds that

$$2m \geq \mathbf{E}(d(A_T)) \geq d(A_0) + \sum_{t=0}^{T-1} \mathbf{E}\left(\frac{|C_t|}{2}\right). \tag{22}$$

We compute a lower bound on $\mathbf{E}(|C_t|)$. By Lemma 4.1, we have that if the event $\{\mathbf{infec}(v) \geq t+1\}$ holds, then $\{|C_t| \geq 1\}$ holds as well. Hence,

$$\mathbf{E}(|C_t|) \geq \mathbf{P}(|C_t| \geq 1) \geq \mathbf{P}(\mathbf{infec}(v) \geq t+1). \tag{23}$$

From (22) and (23), we conclude that for all $T \geq 1$,

$$4m \geq \sum_{t=0}^{T-1} \mathbf{P}(\mathbf{infec}(v) \geq t+1).$$

Hence, by taking T tending to infinity,

$$4m \geq \sum_{t=1}^{\infty} \mathbf{P}(\mathbf{infec}(v) \geq t) = \mathbf{E}(\mathbf{infec}(v)).$$

□

The next corollary proves the bound (9) of Theorem 2.4 and, via duality (Theorem 2.3), the bound (2) of Theorem 2.1.

Corollary 4.4. *There exists a constant $c > 0$ such that*

$$\boldsymbol{P}(\mathbf{infec}(v) > cm \log n) \leq n^{-3}.$$

Proof. By Markov's inequality

$$\mathbf{P}(\mathbf{infec}(v) \geq 2\mathbf{E}(\mathbf{infec}(v))) \leq 1/2.$$

Consider the following algorithm. At time $T = 2\mathbf{E}(\mathbf{infec}(v))$, we check if the process infected the whole graph. If so, we stop, otherwise we drop all the infection and restart the algorithm until it stops. Due to the fact that we infect the graph with probability at least $1/2$ by time $2\mathbf{E}(\mathbf{infec}(v))$, the probability that we iterate the algorithm more than $3\lceil \log_2 n \rceil$ times is at most n^{-3}. The conclusion follows from the fact that $\mathbf{E}(\mathbf{infec}(v)) \leq 4m$, as claimed in Theorem 4.3. □

4.2 Sequential Analysis: Proof of Bound (8) in Theorem 2.4

The proof of bound (8) of Theorem 2.4 requires a more subtle argument. One of the difficulties of studying the BIPS process is that the one-step difference, $d(A_{t+1}) - d(A_t)$ has a large range. For instance, in the complete graph, the whole graph may become infected or healthy (by exception of the source) in one step. This discourages us from attempting to use a raw concentration inequality to prove that $\mathbf{infec}(v)$ is concentrated around its mean (but we note that such concentration, if established, and Theorem 4.3 might give $\mathbf{infec}(v) = \mathcal{O}(m)$ w.h.p.). To face this problem, one option is to study the actual distribution of $d(A_{t+1}) - d(A_t)$, but this seems hard because of the non-monotonic behaviour of A_t. Instead we follow a different approach, using the fact that $d(A_{t+1}) - d(A_t)$, given A_t, can be written as a sum of independent random variables and the range of each of these variables is an interval of length at most $d_{\max}$. If the maximum degree $d_{\max}$ is not too large, then we will be able to conclude that w.h.p. $\mathbf{infec}(v) = \mathcal{O}(m)$.

We proceed to formalise the above idea, starting from Equation (18):

$$d(A_{t+1}) = d(A_t) + \sum_{u \in C_t} (d(u)X_{t+1,u} - d_{A_t}(u)), \tag{24}$$

where $X_{\tau,u}$ indicates whether $u \in A_\tau$. By iteratively applying Equation (24), for any $t \leq \mathbf{infec}(v)$, we have

$$d(A_t) \quad = \quad d(v) + \sum_{\tau=0}^{t-1} \sum_{u \in C_\tau} (d(u)X_{\tau+1,u} - d_{A_\tau}(u)). \tag{25}$$

From Lemma 4.1, it holds that $|C_\tau| > 0$ for all $0 \leq \tau < \mathbf{infec}(v)$, hence, for all $1 \leq t \leq \mathbf{infec}(v)$, we write Equation (25) as

$$d(A_t) \quad = \quad d(v) + \sum_{l=1}^{\nu(t)} Y_l. \tag{26}$$

Here, $\nu(0) = 0$, $\nu(t) \equiv \sum_{\tau=0}^{t-1} |C_\tau|$ for $t \geq 1$, and $Y_{\nu(\tau)+i} \equiv d(u)X_{\tau+1,u} - d_{A_\tau}(u)$, for $0 \leq \tau < t$ and $1 \leq i \leq |C_\tau|$, where u is the i-th smallest vertex of C_τ in some arbitrary but fixed ordering of the set of vertices V. Since $\nu(0) = 0$, and $1 \leq |C_\tau| \leq n$, we have

$$t \;\leq\; \nu(t) \;<\; \nu(t-1) + n. \tag{27}$$

We now refer to one step of BIPS as one *round*, reserving the term *step* for the process of considering the vertices in C_τ one by one. We say that round t, with $1 \leq t \leq \mathbf{infec}(v)$, consists of $|C_{t-1}|$ steps, with the random variable $Y_{\nu(t-1)+i}$ corresponding to the i-th step of this round. Thus, we can view the whole BIPS process as one sequence of single steps $1, 2, \ldots, l, \ldots$, which are grouped into rounds.

Even though the BIPS process finishes at round $\mathbf{infec}(v)$, the sequence $(A_t)_{t\geq 0}$ is defined in the natural way for all $t \geq 0$. For $t \geq \mathbf{infec}(v)$, $A_t = V$ and thus $d(A_t) = 2m$. The sequence (Y_l) is defined for $1 \leq l < \nu(\mathbf{infec}(v))$, that is, until the completion of the BIPS process. For technical convenience, we set $Y_l = 1$ for all $l \geq \nu(\mathbf{infec}(v))$, so the process (Y_l) is defined for all $l \geq 0$. The choice of the value 1 will become clear later.

Observe that the random variables Y_l are not independent. Indeed, the distribution of Y_l depends on the values of the variables Y_i, for $1 \leq i \leq l-1$. For any fixed $l \geq 1$, and an arbitrary sequence of numbers $y_1, y_2, \ldots, y_{l-1}$, we have two possibilities. Either the given sequence of numbers is not a feasible realisation of the sequence of random variables $Y_1, Y_2, \ldots, Y_{l-1}$, or it is feasible realisation and shows in full the evolution of the BIPS process until step $l-1$, determining the distribution of the variable Y_l.

In particular, if $Y_1 = y_1, Y_2 = y_2, \ldots, Y_{l-1} = y_{l-1}$, then by starting from the known initial sets $A_0 = \{v\}$ and C_0, and using the fact that the vertices of C_t are considered according to a fixed ordering of all vertices of V, the values of $Y_1, Y_2, \ldots$ identify the vertices in A_1 (and set A_1 gives set C_1). This can be continued to identify set A_2, then set A_3, and so on. Finally, if the given sequence of the values of variables Y_i's is feasible, then either the process has completed before step l, so $Y_l = 1$, or we identify the round t which includes the step l. In the latter case, we are able to recover the set $A_{t-1} \subsetneq V$ of vertices which infect other vertices during round t, and the vertex u considered in step l. In both cases, we get the distribution of the random variable Y_l.

Equation (26) suggests that instead of analysing the sequence $(d(A_t))_{t\geq 0}$, we can analyse the sequence of sums $R_q = \sum_{l=1}^{q} Y_l$, $q \geq 0$. There is a technical complication here because an element R_q of the sequence $(R_q)_{q\geq 0}$ corresponds to an element of the sequence $(d(A_t))_{t\geq 0}$ only if $q = \nu(t)$ for some $t \geq 0$. This means that a large value of some R_q does not immediately guarantee a large value of $d(A_t)$. We will be looking therefore for a long sub-sequence $R_q, R_{q+1}, \ldots, R_{q'}$ of large values which will guarantee that one of the indices $\{q, q+1 \ldots, q'\}$ corresponds to a value $\nu(t)$ for some t, implying a large value of some $d(A_t)$.

More precisely, we have the following relationship between the sequences $(d(A_t))_{t\geq 0}$ and $(R_q)_{q\geq 0}$. For each $1 \leq k$ and each $t \geq 1$,

$$\{(A_t \neq V) \wedge (d(A_t) < d(v) + k)\} \subseteq \{\exists\ q : t \leq q < tn\ \wedge\ R_q < k\}. \tag{28}$$

The above holds for the following reason. Consider an execution of the BIPS process such that $A_t \neq V$ and $d(A_t) < d(v) + k$. From (26), $R_{\nu(t)} = \sum_{l=1}^{\nu(t)} Y_l < k$, and from (27), $t \leq \nu(t) < tn$. Thus $R_q < k$, for some $t \leq q < tn$.

We proceed to derive a lower bound on the conditional expectation of Y_l, $l \geq 1$, given the values of the variables $Y_1, Y_2, \ldots, Y_{l-1}$. If these values show that the BIPS process has already infected the whole graph (that is, $l > \nu(T)$, where $T = \mathbf{infec}(v)$), then $Y_l = 1$ and $\mathbf{E}(Y_l | Y_1, Y_2, \ldots, Y_{l-1}) = 1$. Otherwise, let u denote the vertex corresponding to Y_l (that is, the vertex considered in step l), let t denote the index of the current round (that is, the round which includes step l), and let $A = A_{t-1}$ and $C = C_{t-1}$. As mentioned above, u, t and A_{t-1} (and thus C_{t-1}) are fully defined by the values of variables $Y_1, Y_2, \ldots, Y_{l-1}$. If u is the source v, then $v \in C$, so $d_A(v) \leq d(v) - 1$ and $Y_l = d(v) - d_A(v) \geq 1$. If $u \neq v$, then (see (20)))

$$\begin{aligned} \mathbf{E}(Y_l | Y_1, Y_2, \ldots, Y_{l-1}) &= d(u)\left(1 - \left(1 - \frac{d_A(u)}{d(u)}\right)^2\right) - d_A(u) \\ &\geq 1 - \frac{1}{d(u)} \geq \frac{1}{2}. \end{aligned} \tag{29}$$

The inequalities (explained in Equation (20)) hold because $u \in C$. Therefore, in all cases, we have

$$\mathbf{E}(Y_l | Y_1, Y_2, \ldots, Y_{l-1}) \geq \frac{1}{2}. \tag{30}$$

Lemma 4.5. *For any $q \geq 1$ and $\delta \geq 0$, it holds that*

$$\boldsymbol{P}\Big(R_q \leq \frac{q}{2} - \delta\Big) \;=\; \boldsymbol{P}\left(\sum_{l=1}^{q} Y_l \leq \frac{q}{2} - \delta\right) \;\leq\; \exp\left(-\frac{\delta^2}{8qd_{\max}^2}\right).$$

Proof. Consider random variables $Z_l = Y_l - \mathbf{E}(Y_l|Y_1, \ldots, Y_{l-1})$, for $l \geq 1$. From the definition of the random variables Y_l, we have that $|Z_l| \leq 2d_{\max}$. Also, it holds that

$$\mathbf{E}(Z_l|Y_1, \ldots, Y_{l-1}) = 0. \tag{31}$$

Therefore, the sequence $S_q = \sum_{l=1}^{q} Z_l$, for $q \geq 1$, is a martingale with respect to the sequence Y_l, $l \geq 1$. Also, for each $q \geq 1$, it holds that $|S_q - S_{q-1}| = |Z_q| \leq 2d_{\max}$. From Azuma's inequality (Theorem A.2 in Appendix), it holds that for all $\delta \geq 0$,

$$\mathbf{P}(S_q < -\delta) \;\leq\; \exp\left(\frac{-\delta^2}{8qd_{\max}^2}\right). \tag{32}$$

Thus with probability at least $1 - \exp(-\delta^2/(8qd_{\max}^2))$, we have

$$\sum_{l=1}^{q} Y_l \;=\; S_q + \sum_{l=1}^{q} \mathbf{E}(Y_l|Y_1, \ldots, Y_{l-1}) \;\geq\; S_q + \frac{q}{2} \;\geq\; \frac{q}{2} - \delta. \tag{33}$$

The first inequality of (33) comes from (30). □

The proof of bound (8) of Theorem 2.4 follows from the next lemma by choosing $k = 2m - d(v)$ and $C = 3$. Observe that the event that $d(A_t) < 2m$ is the same as the event that $\mathbf{infec}(v) > t$.

Lemma 4.6. *Consider the BIPS process on a connected graph with n vertices, m edges and the maximum vertex degree $d_{\max}$. For any constant $C > 0$, define $C' = 64(C + 2)$. Then, for any $1 \leq k \leq 2m - d(v)$, and $t(k) = 4k + C'(d_{\max})^2 \log n$,*

$$\boldsymbol{P}(\exists\, t \geq t(k) :\; d(A_t) < d(v) + k) \;=\; \mathcal{O}(n^{-C}).$$

Proof. From Equation (28) and Lemma (4.5), it holds

$$\begin{aligned}
\mathbf{P}(\exists\; t \geq t(k) : d(A_t) < d(v) + k) \;&\leq\; \mathbf{P}\left(\bigcup_{t \geq t(k)} \{\exists\; q : t \leq q \leq tn \wedge\; R_q < k\}\right) \\
&\leq\; \mathbf{P}\left(\bigcup_{q \geq t(k)} \{R_q < k\}\right) \\
&\leq\; \sum_{q=t(k)}^{\infty} \exp\left(-\frac{(q/2 - k)^2}{8qd_{\max}^2}\right) \\
&\leq\; \sum_{q=t(k)}^{\infty} \exp\left(-\frac{q/2 - k}{32d_{\max}^2}\right) \qquad (34) \\
&\leq\; \sum_{j=0}^{\infty} \exp\left(-\frac{(C'/2)d_{\max}^2 \log(n) + j/2}{32d_{\max}^2}\right) \\
&=\; \mathcal{O}(d_{\max}^2) \cdot \exp\left(-\frac{C' \log(n)}{64}\right) \\
&=\; \mathcal{O}(d_{\max}^2) \cdot n^{-(C+2)} = \mathcal{O}(n^{-C}). \qquad (35)
\end{aligned}$$

The inequality in (34) holds because $(q/2 - k)/q \geq 1/4$ for all $q \geq t(k) \geq 4k$. □

5 The BIPS Process on Regular Graphs

In the analysis of the BIPS process on regular graphs, we track the size of the current infection set rather than the degree of this set. This analysis is done in two phases. The first phase deals with small infection sizes, while the second phase considers large infections. We begin by analysing the second phase since the arguments are easier.

5.1 Large Infection Size

We begin our analysis by giving a lower bound of the size of A_{t+1} given A_t.

Lemma 5.1. *Let G be a connected r-regular graph on n vertices, with $\lambda < 1$ where λ is the second absolute eigenvalue of the random-walk transition matrix. Let A_t be the infected set after round t of the BIPS process with $k = 2$, then*

$$\mathbf{E}\big(|A_{t+1}| \,\big|\, A_t = A\big) \;\geq\; |A|(1 + (1-\lambda^2)(1 - |A|/n)). \tag{36}$$

Proof. A direct computation gives us

$$\begin{aligned}
\mathbf{E}\big(|A_{t+1}| \,\big|\, A_t = A\big) &= 1 + \sum_{u \in V\setminus\{v\}} \mathbf{P}(X_{t+1,u} = 1 \mid A_t = A) \\
&= 1 + \sum_{u \in V\setminus\{v\}} \big(1 - (1 - d_A(u)/r)^2\big) \\
&\geq \sum_{u \in V} \big(1 - (1 - d_A(u)/r)^2\big) \\
&= \sum_{u \in V} \left(2\frac{d_A(u)}{r} - \frac{d_A(u)^2}{r^2}\right) \\
&= 2|A| - \sum_{u \in V} \frac{d_A(u)^2}{r^2}.
\end{aligned} \tag{37}$$

In the last step, we use the fact that $\sum_{u\in V} d_A(u) = \sum_{u\in A} d(u) = r|A|$. Let $P = P(G)$ be the transition matrix of a simple random walk on G. Let $P(x, A) = \sum_{y\in A} P(x, y) = d_A(x)/r$. From (37), we have

$$\mathbf{E}\big(|A_{t+1}| \,\big|\, A_t = A\big) \;\geq\; 2A - \sum_{x\in V} (P(x, A))^2. \tag{38}$$

Observe that $\sum_{x\in V}(P(x, A))^2 = \langle P\mathbb{1}_A, P\mathbb{1}_A\rangle = \|P\mathbb{1}_A\|^2$, where $\mathbb{1}_A$ is the characteristic vector of A and $P\mathbb{1}_A$ represents the standard matrix-vector product. As P is symmetric, it has an orthonormal basis of right eigenvectors $f_1, ..., f_n$, i.e., $\|f_i\| = 1$ and $\langle f_i, f_j\rangle = 0$ for $i \neq j$. For any real n-dimensional vector g, we have $g = \sum_{i=1}^n \langle g, f_i\rangle f_i$ and $\|g\|^2 = \sum_{i=1}^n \langle g, f_i\rangle^2$. Here $f_1 = (1/\sqrt{n}, 1/\sqrt{n}, \ldots, 1/\sqrt{n})$ is the unique eigenvector with eigenvalue 1, and $\langle \mathbb{1}_A, f_1\rangle = A/\sqrt{n}$. Thus,

$$\begin{aligned}
\|P\mathbb{1}_A\|^2 &= \|P\sum_{i=1}^n \langle \mathbb{1}_A, f_i\rangle f_i\|^2 \;=\; \|\sum_{i=1}^n \langle \mathbb{1}_A, f_i\rangle P f_i\|^2 \\
&= \|\sum_{i=1}^n \langle \mathbb{1}_A, f_i\rangle \lambda_i f_i\|^2 \;=\; \sum_{i=1}^n \langle \mathbb{1}_A, f_i\rangle^2 \lambda_i^2 \|f_i\|^2 \\
&\leq (1-\lambda^2)\langle \mathbb{1}_A, f_1\rangle^2 + \lambda^2 \sum_{i=1}^n \langle \mathbb{1}_A, f_i\rangle^2 \\
&= (1-\lambda^2)\frac{|A|^2}{n} + \lambda^2 \|\mathbb{1}_A\|^2 \\
&= (1-\lambda^2)\frac{|A|^2}{n} + \lambda^2 |A|.
\end{aligned} \tag{39}$$

Hence, Inequalities (38) and (39) imply

$$\mathbf{E}\big(|A_{t+1}|\,\big|\,A_t = A\big) \;\geq\; 2|A| - \lambda^2|A| - (1-\lambda^2)\frac{|A|^2}{n},$$

which is equivalent to (36). □

A direct application of the lemma above allows us to analyse the second phase of the process which begins when the number of infected vertices is greater than $K\log n/(1-\lambda)^2$, for a large enough constant $K > 0$. The first lemma below considers the case when the number of the infected vertices is between $K\log(n)/(1-\lambda)^2$ and $9n/10$, and the second lemma deals with the case when the number of infected vertices is at least $9n/10$.

Lemma 5.2. *There exists a sufficiently large constant $K > 0$, such that for each sufficiently large n, each connected regular n-vertex graph G, and each round $t > 0$ of the BIPS process on G, if $1-\lambda \geq \sqrt{2K(\log n)/n}$ and $|A_t| \geq K\log n/(1-\lambda)^2$, then with probability at least $1-n^{-4}$, at least $(9/10)$-th of the whole graph is infected within $\mathcal{O}(\log n/(1-\lambda))$ subsequent rounds.*

Proof. Assume A_t has size less than $9n/10$ but at least $K\log n/(1-\lambda)^2$. Observe that the lower bound on $1-\lambda$ implies that $K\log n/(1-\lambda)^2 \leq n/2$. Then from Lemma 5.1,

$$\begin{aligned}\mathbf{E}\big(|A_{t+1}|\,\big|\,A_t\big) &\geq |A_t|(1+(1-\lambda^2)(1-9/10))\\ &\geq |A_t|\left(1+\frac{1-\lambda}{10}\right).\end{aligned}$$

Let $\varepsilon = \sqrt{10\log n/|A_t|}$ and, for a given set A_t, view the size of A_{t+1} as the sum of n independent Bernoulli random variables (with potentially different probabilities of success). By using the Chernoff bound for the lower tail of the sum of Bernoulli random variables (Theorem A.1 in Appendix), we get

$$\begin{aligned}\mathbf{P}\Big(|A_{t+1}| < (1-\varepsilon)\mathbf{E}\big(|A_{t+1}|\,\big|\,A_t\big)\,\Big|\,A_t\Big) &\leq \exp\big\{-\varepsilon^2\mathbf{E}\big(|A_{t+1}|\,\big|\,A_t\big)/2\big\}\\ &= \exp\left\{-\frac{5\log n}{|A_t|}\mathbf{E}\big(|A_{t+1}|\,\big|\,A_t\big)\right\}\\ &\leq \exp\left\{-5\log n\left(1+\frac{1-\lambda}{10}\right)\right\} \;\leq\; \frac{1}{n^5}.\end{aligned}$$

Choosing $K = 4000$, we have, by hypothesis $|A_t| \geq K\log n/(1-\lambda)^2$, that $\varepsilon \leq (1-\lambda)/20$. Therefore, with probability at least $1-n^{-5}$ we have

$$|A_{t+1}| \;\geq\; (1-\varepsilon)\mathbf{E}\big(|A_{t+1}|\,\big|\,A_t\big) \;\geq\; |A_t|\left(1+\frac{1-\lambda}{10}\right)\left(1-\frac{1-\lambda}{20}\right) \;\geq\; |A_t|\left(1+\frac{1-\lambda}{23}\right).$$

Hence, after $23/(1-\lambda)$ rounds, the size of infection has at least doubled with probability at least $1-23n^{-5}/(1-\lambda)$. Finally, with probability at least $1-23(\log_2 n)n^{-5}/(1-\lambda) \geq 1-n^{-4}$, after $\mathcal{O}(\log n/(1-\lambda))$ rounds, the infection covers at least $9n/10$ vertices. □

Lemma 5.3. *Let G be a connected regular n-vertex graph with $1-\lambda \geq c\sqrt{\log n/n}$, for a suitably large constant c. If the BIPS process starts with at least $(9/10)n$ infected vertices, then with probability $1-O(n^{-5})$ the whole graph is infected within $7\log n/(1-\lambda)$ rounds.*

Proof. For convenience, let A_0 and B_0 be the infected and non-infected sets at the beginning of this phase, and denote $q = 9/10$. We assume that $|A_0| \geq qn$. Let A_t and $B_t = V \setminus A_t$ be the infected and non-infected sets after t rounds, respectively. From (36), we get the following bound on the expected size of the set A_{t+1} given the set A_t.

$$\mathbf{E}\big(|A_{t+1}|\,\big|\,A_t\big) \;\geq\; |A_t| + (n-|A_t|)(1-\lambda^2)|A_t|/n. \tag{40}$$

The corresponding inequality for the expected size of the set $B_{t+1} = V \setminus A_{t+1}$ is

$$\begin{aligned}\mathbf{E}\big(|B_{t+1}|\,\big|\,B_t\big) &\leq |B_t| - |B_t|(1-\lambda^2)|A_t|/n\\ &= |B_t|(1-(1-\lambda^2)|A_t|/n).\end{aligned} \tag{41}$$

Let $|A_t| = k$. By applying the law of total probability and Equation (41), we get

$$\begin{aligned}
\mathbf{E}(|B_{t+1}|) &= \sum_{k=qn}^{n} \mathbf{E}\big(|B_{t+1}| \,\big|\, |B_t| = n-k\big) \cdot \mathbf{P}(|B_t| = n-k) \;+\; \mathbf{E}\big(|B_{t+1}| \,\big|\, |A_t| < qn\big) \cdot \mathbf{P}(|A_t| < qn) \\
&\le \sum_{k=qn}^{n} (n-k)(1-(1-\lambda^2)k/n) \cdot \mathbf{P}(|B_t| = n-k) \;+\; n\mathbf{P}(|A_t| < qn) \\
&\le (1-(1-\lambda^2)q) \sum_{k=qn}^{n} (n-k)\mathbf{P}(|B_t| = n-k) \;+\; n\mathbf{P}(|A_t| < qn) \\
&\le (1-(1-\lambda^2)q) \cdot \mathbf{E}(|B_t|) + n\mathbf{P}(|A_t| < qn). \qquad (42)
\end{aligned}$$

We next prove that

$$\mathbf{P}(|A_t| < qn) \le tn^{-8}. \qquad (43)$$

To check that the above inequality holds, consider the event $\mathcal{Q}_t = \{|A_i| \ge qn, i = 0, \dots, t\}$. We are going to prove that $\mathcal{Q}_t$ holds with high probability. Indeed

$$\begin{aligned}
\mathbf{P}(\mathcal{Q}_t) &= \mathbf{P}(\mathcal{Q}_t|\mathcal{Q}_{t-1})\mathbf{P}(\mathcal{Q}_{t-1}) + \mathbf{P}(\mathcal{Q}_t|\mathcal{Q}_{t-1}^c)\mathbf{P}(\mathcal{Q}_{t-1}^c) \\
&\ge \mathbf{P}(\mathcal{Q}_t|\mathcal{Q}_{t-1})\mathbf{P}(\mathcal{Q}_{t-1}). \qquad (44)
\end{aligned}$$

Observe that A_t depends only on A_{t-1}, so

$$\mathbf{P}(\mathcal{Q}_t|\mathcal{Q}_{t-1}) \;=\; \mathbf{P}\big(|A_t| \ge qn \,\big|\, |A_{t-1}| \ge qn\big), \qquad (45)$$

and by a standard coupling argument

$$\mathbf{P}\big(|A_t| \ge qn \,\big|\, |A_{t-1}| \ge qn\big) \;\ge\; \mathbf{P}\big(|A_t| \ge qn \,\big|\, |A_{t-1}| = \lfloor qn \rfloor\big). \qquad (46)$$

Choose $\varepsilon = \sqrt{16 \log n / \lfloor qn \rfloor}$, then by the Chernoff bound

$$\begin{aligned}
\mathbf{P}\big(|A_t| < (1-\varepsilon)\mathbf{E}\big(A_t \,\big|\, |A_{t-1}| = \lfloor qn \rfloor\big) \,\big|\, |A_{t-1}| = \lfloor qn \rfloor\big) &\le \exp\big\{-\varepsilon^2 \mathbf{E}\big(|A_t| \,\big|\, |A_{t-1}| = \lfloor qn \rfloor\big)/2\big\} \\
&\le \exp\big\{-\varepsilon^2 \lfloor qn \rfloor/2\big\} \;=\; \frac{1}{n^8}.
\end{aligned}$$

Since we assume that $1 - \lambda \ge c\sqrt{\log n / n}$, for a suitably large constant c, we have $\varepsilon \le (1-\lambda)/15$ for all $n \ge n_0(c)$. Thus, if $|A_{t-1}| = \lfloor qn \rfloor$, then with probability at least $1 - n^{-8}$,

$$\begin{aligned}
|A_t| &\ge (1-\varepsilon)\mathbf{E}\big(|A_t| \,\big|\, |A_{t-1}| = \lfloor qn \rfloor\big) \\
&\ge \lfloor qn \rfloor \left(1 + \frac{1-\lambda}{10}\right)\left(1 - \frac{1-\lambda}{15}\right) \;\ge\; qn,
\end{aligned}$$

where the second inequality follows from Lemma 5.1 and the third one holds for all sufficiently large n. This, together with (44), (45) and (46), gives

$$\mathbf{P}(\mathcal{Q}_t) \ge (1-n^{-8})^t \ge 1 - tn^{-8}.$$

Observe that $\mathbf{P}(|A_t| \ge qn) \ge \mathbf{P}(\mathcal{Q}_t) \ge 1 - tn^{-8}$, so Equation (43) holds.

Let us return to (42). By using Inequality (43), we have

$$\mathbf{E}(|B_{t+1}|) \;\le\; (1-(1-\lambda^2)q)\mathbf{E}(|B_t|) + tn^{-7}. \qquad (47)$$

Denote $\theta = 1 - (1-\lambda^2)q$, then by iterating (47) and by using that $|B_0| \le (1-q)n$, we get

$$\mathbf{E}(|B_t|) \;\le\; \theta^t(1-q)n + \mathcal{O}(t^2 n^{-7}) \le n\theta^t + \mathcal{O}(t^2 n^{-7}).$$

Choosing $T = 6\log n/\log(1/\theta)$, and applying Markov's inequality, we get

$$\begin{aligned}\mathbf{P}(|B_T| \geq 1) &\leq \mathbf{E}(|B_T|) \leq n\theta^T + O(T^2 n^{-7}) \\ &= n^{-5} + O(T^2 n^{-7}). \end{aligned} \tag{48}$$

Observe that for $0 < \theta < 1$, we have $(1-\theta) \leq \log(1/\theta)$, so

$$\begin{aligned} T &= 6\log n/(\log(1/\theta)) \leq 6\log n/(1-\theta) \\ &= 6\log n/(q(1-\lambda^2)) \leq 6\log n/(q(1-\lambda)) \\ &\leq 7\log n/(1-\lambda) = \mathcal{O}(n). \end{aligned}$$

The final $\mathcal{O}(n)$ bound on T (which we need to bound T on the right-hand side in (48)) follows from the assumption that $1-\lambda \geq c\sqrt{(\log n)/n}$. Thus (48) implies that $\mathbf{P}(B_T \geq 1) = \mathcal{O}(n^{-5})$, so the probability $\mathbf{P}(B_T = 0)$, that is, the probability that all vertices are infected at time T, is $1 - \mathcal{O}(n^{-5})$ and $T \leq 7\log n/(1-\lambda)$. □

In order to apply the two lemmas above together, we do the following. Start the BIPS process with $|A_0| \geq K\log n/(1-\lambda)$ where K is the constant from Lemma 5.2. Let T be the first time such that $|A_T| \geq 9n/10$. From Lemma 5.2, we know that there exists a constant C, such that $T \leq C\log n/(1-\lambda)$ with probability at least $1 - n^{-4}$.

Let $\mathcal{H} = \{A \subseteq V : |A| \geq 9n/10\}$ and let $t = 7\log n/(1-\lambda)$, then

$$\begin{aligned}\mathbf{P}\big(A_{T+t} = V \,\big|\, |A_T| \geq 9n/10\big) &= \sum_{A\in\mathcal{H}} \mathbf{P}(A_{T+t} = V \mid A_T = A)\cdot\mathbf{P}(A_T = A) \\ &= \sum_{A\in\mathcal{H}} \mathbf{P}(A_t = V \mid A_0 = A)\cdot\mathbf{P}(A_T = A) \qquad (49) \\ &\geq \sum_{A\in\mathcal{H}} (1-\mathcal{O}(n^{-5}))\cdot\mathbf{P}(A_T = A) = 1 - \mathcal{O}(n^{-5}). \qquad (50)\end{aligned}$$

Equation (49) holds as a consequence of the strong Markov property. Inequality (50) follows from Lemma 5.3. We conclude that with probability at least $1 - 2n^{-4}$ the process finishes within $T + t = \mathcal{O}(\log n/(1-\lambda))$ steps.

Corollary 5.4. *Suppose that we start a BIPS process with infection size $|A_0| \geq K\log n/(1-\lambda)^2$ and assume that $1-\lambda \geq c\sqrt{\log n/n}$, where K and c are large enough constants and $c \geq \sqrt{2K}$. Then the process infects the whole graph in $\mathcal{O}(\log n/(1-\lambda))$ rounds with probability at least $1 - 2n^{-4}$.*

5.2 Small Infection Sizes

As seen in Corollary 5.4, when the size of the infection is fairly large, the whole graph is infected in $\mathcal{O}(\log n/(1-\lambda))$ rounds. In this section, we prove that in $\mathcal{O}(\min\{1/(1-\lambda)^2,\, r/(1-\lambda) + r^2\}\log n)$ rounds the infection reaches the necessary size to apply Corollary 5.4, which leads us to the proof of Theorem 2.5.

The main difference between the analysis of the small and large infection cases, is that we can use the concentration behaviour of large infections to track the size of A_t and thus prove that it increases in each round by a substantial amount. This approach, however, does not work for the initial part of the BIPS process, when the sizes of sets A_t are small. Indeed, when the size of the infection is relatively small, in one step the infection can either grow or shrink depending on several factors, including randomness, graph structure and the location of infected vertices.

In this section, we provide the analysis of the early stages of the BIPS process which will lead us to the proof of Theorem 2.5. The main part of the analysis is given in Lemma 5.5, which, for parameters $1 \leq \alpha < \beta$, relates the probability of the infection reaching size β by round $t + T(\alpha,\beta)$ to the probability of reaching size α by round t. Lemma 5.6, which is a simple consequence of Lemma 5.5, gives us the proof of bound (11) of Theorem 2.5. Finally, Lemma 5.8, which uses Lemma 5.5 and Lemma 4.6, gives us the proof of bound (10) of Theorem 2.5.

We begin our analysis by introducing the necessary notation. Define the quantity $\Delta = K\log(n)/(1-\lambda)^2$ where K is the constant from Lemma 5.2. Note that once the infection has size $|A_t| \geq \Delta$, it is sufficiently large to apply the results of Section 5.1. We therefore refer to Δ as the final target size of infection. Let $T \geq t \geq 0$ be two integers, and let a parameter α, where $1 \leq \alpha \leq \Delta$, be an intermediate target size of infection. We define the following events $\mathcal{A}_{t,\alpha}$, $\mathcal{E}^T_{t,\alpha}$ and $\mathcal{E}_{t,\alpha}$.

$$\begin{aligned}
\mathcal{A}_{t,\alpha} &= \{|A_t| \geq \alpha\} \cup \bigcup_{i=0}^{t-1} \{|A_i| \geq \Delta\}, \\
\mathcal{E}^T_{t,\alpha} &= \bigcap_{i=t}^{T} \mathcal{A}_{i,\alpha}, \\
\mathcal{E}_{t,\alpha} &= \bigcap_{i=t}^{\infty} \mathcal{A}_{i,\alpha}.
\end{aligned}$$

The event $\mathcal{A}_{t,\alpha}$ says that at round t the infection size is at least α (the intermediate target) or it has already hit the final target of Δ at some earlier round. The event $\mathcal{E}_{t,\alpha}$ says that at round t the infection size is at least α and will not drop below α before reaching the final target Δ, or the infection size has already reached Δ before round t. Observe that due to the source vertex, it holds that $|A_t| \geq 1$ for all t, therefore $\mathbf{P}(\mathcal{E}_{0,1}) = 1$. Observe also that the event $\mathcal{E}_{t,\alpha}$ implies that there exists a round $s \leq t$ such that $|A_s| \geq \alpha$.

In Lemma 5.5 below we have two intermediate targets α and β for the infection size, where $1 \leq \alpha < \beta \leq \Delta$. The lemma says (roughly) that for some appropriately large $T = T(\alpha, \beta, n, \lambda)$, if the infection size is at least α at some round t, then w.h.p. it will reach the second threshold of β within the subsequent T rounds.

Lemma 5.5. *Let G be a connected n-vertex regular graph with $1-\lambda \geq \sqrt{2K(\log n)/n}$, where K is the constant from Lemma 5.2 (implying that $\Delta \leq n/2$). Let α and β be such that $1 \leq \alpha < \beta \leq \Delta$, and let $t \geq 0$ be a non-negative integer. Choosing $T = \frac{6}{\alpha}\left(\frac{\beta}{1-\lambda} + \frac{2C\log n}{(1-\lambda)^2}\right)$, where $C \geq 1$, for large enough n we have*

$$\boldsymbol{P}(\mathcal{E}_{t+T,\beta}) \geq \boldsymbol{P}(\mathcal{E}_{t,\alpha}) - 2n^{-C}. \tag{51}$$

Proof. We derive an upper bound on the probability of $\mathcal{E}^c_{t+T,\beta}$, the complement event of $\mathcal{E}_{t+T,\beta}$. We have

$$\begin{aligned}
\mathbf{P}(\mathcal{E}^c_{t+T,\beta}) &\leq \mathbf{P}(\mathcal{E}^c_{t+T,\beta} \cap \mathcal{E}_{t,\alpha}) + \mathbf{P}(\mathcal{E}^c_{t,\alpha}) \\
&\leq \sum_{s=t+T}^{\infty} \mathbf{P}(\mathcal{A}^c_{s,\beta} \cap \mathcal{E}_{t,\alpha}) + \mathbf{P}(\mathcal{E}^c_{t,\alpha}). \qquad (52)
\end{aligned}$$

We focus on $\mathbf{P}(\mathcal{A}^c_{s,\beta} \cap \mathcal{E}_{t,\alpha})$ for $s \geq t+T$. From the definition of $\mathcal{A}_{t,\alpha}$,

$$\mathcal{A}^c_{s,\beta} = \{|A_s| < \beta\} \cap \bigcap_{i=0}^{s-1} \{|A_i| < \Delta\}. \tag{53}$$

For $s > t$, define the event $\mathcal{B}_{s,t,\alpha} = \bigcap_{i=t}^{s} \{\Delta > |A_i| \geq \alpha\}$. Then

$$\begin{aligned}
\bigcap_{i=0}^{s-1} \{|A_i| < \Delta\} \cap \mathcal{E}_{t,\alpha} &\subseteq \bigcap_{i=0}^{s-1} \{|A_i| < \Delta\} \cap \mathcal{E}^{s-1}_{t,\alpha} \qquad (54) \\
&= \left(\bigcap_{i=0}^{s-1} \{|A_i| < \Delta\}\right) \cap \left(\bigcap_{i=t}^{s-1} \left(\{|A_i| \geq \alpha\} \cup \bigcup_{j=0}^{i-1} \{|A_j| \geq \Delta\}\right)\right) \qquad (55) \\
&\subseteq \bigcap_{i=t}^{s-1} \{\Delta > |A_i| \geq \alpha\} = \mathcal{B}_{s-1,t,\alpha}. \qquad (56)
\end{aligned}$$

The inclusion relation (54) holds because by definition $\mathcal{E}_{t,\alpha} \subseteq \mathcal{E}^{s-1}_{t,\alpha}$. The Equation (55) is obtained by substituting the definition of $\mathcal{E}^{s-1}_{t,\alpha}$, and the final inclusion relation is easy to verify. From (53) and (56), we

have

$$\mathcal{A}^c_{s,\beta} \cap \mathcal{E}_{t,\alpha} = \{|A_s| < \beta\} \cap \bigcap_{i=0}^{s-1} \{|A_i| < \Delta\} \cap \mathcal{E}_{t,\alpha} \ \subseteq \ \{|A_s| < \beta\} \cap \mathcal{B}_{s-1,t,\alpha}.$$

Let $\phi > 0$ be a parameter to be specified later. We have

$$\begin{aligned} \mathbf{P}(\mathcal{A}^c_{s,\beta} \cap \mathcal{E}_{t,\alpha}) &\leq \mathbf{P}(\{|A_s| < \beta\} \cap \mathcal{B}_{s-1,t,\alpha}) \\ &= \mathbf{P}(e^{-\phi|A_s|}\mathbb{1}_{\mathcal{B}_{s-1,t,\alpha}} \geq e^{-\phi\beta}) \\ &\leq \mathbf{E}(e^{-\phi|A_s|}\mathbb{1}_{\mathcal{B}_{s-1,t,\alpha}})e^{\phi\beta}. \end{aligned} \tag{57}$$

To simplify notation, we write $\mathcal{B}_s$ instead of $\mathcal{B}_{s,t,\alpha}$. We focus on getting a good estimate of $\mathbf{E}(e^{-\phi|A_s|}\mathbb{1}_{\mathcal{B}_{s-1}})$. Denoting $G(s) = \mathbf{E}(e^{-\phi|A_s|}\mathbb{1}_{\mathcal{B}_{s-1}})$ and the sigma algebra $\mathcal{F}_s = \sigma(A_0, \ldots, A_s)$, we have for $s > t$ (using the fact that $\mathcal{B}_{s-1}$ is fixed for given $A_0, A_1, \ldots, A_{s-1}$),

$$\begin{aligned} G(s) &= \mathbf{E}\Big(\mathbf{E}\Big(e^{-\phi|A_s|}\mathbb{1}_{\mathcal{B}_{s-1}} \,\Big|\, \mathcal{F}_{s-1}\Big)\Big) \\ &\leq \mathbf{E}\Big(\mathbb{1}_{\mathcal{B}_{s-1}}\mathbf{E}\Big(e^{-\phi|A_s|} \,\Big|\, \mathcal{F}_{s-1}\Big)\Big). \end{aligned} \tag{58}$$

Remember that given A_{s-1}, the event that $u \in V$ belongs to A_s is independent of other vertices, thus

$$\begin{aligned} \mathbf{E}\Big(e^{-\phi|A_s|}|\mathcal{F}_{s-1}\Big) &= \prod_{u\in V} \mathbf{E}\big(e^{-\phi\mathbb{1}_{\{u\in A_s\}}} \,\big|\, \mathcal{F}_{s-1}\big) \\ &= \prod_{u\in V} \big(e^{-\phi}\mathbf{P}(u \in A_s|A_{s-1}) + \mathbf{P}(u \notin A_s|A_{s-1})\big) \\ &= \prod_{u\in V} \big(1 - (1-e^{-\phi})\mathbf{P}(u \in A_s|A_{s-1})\big) \\ &\leq \prod_{u\in V} \exp\big(-(1-e^{-\phi})\mathbf{P}(u \in A_s|A_{s-1})\big) \\ &= \exp\big(-(1-e^{-\phi})\mathbf{E}\big(|A_s| \,\big|\, A_{s-1}\big)\big) \\ &\leq \exp\big(-(1-e^{-\phi})|A_{s-1}|(1+(1-\lambda^2)(1-|A_{s-1}|/n))\big). \end{aligned}$$

The last inequality follows from Lemma 5.1. Using the above inequality in Inequality (58), we obtain $G(s) \leq \mathbf{E}(\mathbb{1}_{\mathcal{B}_{s-2}}h(A_{s-1}))$, where

$$\begin{aligned} h(A) &= \mathbb{1}_{\{\Delta\geq|A|\geq\alpha\}} \exp\big(-|A|(1-e^{-\phi})(1+(1-\lambda^2)(1-|A|/n))\big) \\ &\leq \mathbb{1}_{\{\Delta\geq|A|\geq\alpha\}} \exp\big(-|A|(1-e^{-\phi})(1+(1-\lambda^2)(1-\Delta/n))\big) \\ &\leq \mathbb{1}_{\{\Delta\geq|A|\geq\alpha\}} \exp\big(-|A|(1-e^{-\phi})(1+(1-\lambda^2)/2)\big). \end{aligned} \tag{59}$$

The last step holds because our assumption that $1 - \lambda \geq \sqrt{2K(\log n)/n}$ implies $\Delta \leq n/2$, so $1 - \Delta/n \geq 1/2$. We write $\delta = 1 + (1-\lambda^2)/2$ and take $\phi = \log\delta > 0$, and continue (59).

$$\begin{aligned} h(A) &\leq \mathbb{1}_{\{\Delta\geq|A|\geq\alpha\}} \exp(-|A|(1-e^{-\phi})\delta) \\ &= \mathbb{1}_{\{\Delta\geq|A|\geq\alpha\}} \exp(-|A|((1-e^{-\phi})\delta - \phi))\exp(-\phi|A|) \\ &\leq \mathbb{1}_{\{\Delta\geq|A|\geq\alpha\}} \exp(-\alpha((1-e^{-\phi})\delta - \phi))\exp(-\phi|A|). \end{aligned} \tag{60}$$

The last inequality is due to the fact that $(1-e^{-\phi})\delta - \phi = e^{\phi} - 1 - \phi \geq 0$. Therefore

$$\begin{aligned} G(s) &\leq \mathbf{E}(\mathbb{1}_{\mathcal{B}_{s-2}}h(A_{s-1})) \\ &\leq \exp(-\alpha((1-e^{-\phi})\delta - \phi))\mathbf{E}(e^{-\phi|A_{s-1}|}\mathbb{1}_{\mathcal{B}_{s-2}}\mathbb{1}_{\{\Delta\geq|A_{s-1}|\geq\alpha\}}) \\ &\leq \exp(-\alpha((1-e^{-\phi})\delta - \phi))G(s-1). \end{aligned} \tag{61}$$

Apply (61) for $s, s-1, \ldots, t+1$ and use the fact that $G(t) \leq 1$ to obtain

$$G(s) \;\leq\; \exp(-(s-t)\alpha((1-e^{-\phi})\delta - \phi)). \tag{62}$$

Thus from (57) and (62), and denoting $\gamma = (1-e^{-\phi})\delta - \phi$, we have

$$\begin{aligned} \mathbf{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) &\leq \exp(-(s-t)\alpha((1-e^{-\phi})\delta - \phi) + \phi\beta) \\ &= \exp(-(s-t)\alpha\gamma + \phi\beta). \end{aligned} \tag{63}$$

Returning to (52) and using (63), we have

$$\begin{aligned} \sum_{s=t+T}^{\infty} \mathbf{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) &\leq \sum_{i=0}^{\infty} e^{-(i+T)\alpha\gamma+\phi\beta} \\ &= e^{-\alpha\gamma T+\phi\beta} \sum_{i=0}^{\infty} e^{-\alpha\gamma i} \;=\; \frac{e^{-\alpha\gamma T+\phi\beta}}{1-e^{-\alpha\gamma T}}. \end{aligned} \tag{64}$$

We proceed to find an upper bound on $-\alpha\gamma T + \phi\beta$. By using the definitions of δ, ϕ and γ in terms of $x = (1-\lambda^2)/2$, we get

$$\begin{aligned} -\alpha\gamma T + \phi\beta &= -\alpha T((1-\delta^{-1})\delta - \log\delta) + \beta\log\delta \\ &= -\alpha T(\delta - 1 - \log\delta) + \beta\log\delta \\ &= -\alpha T(x - \log(1+x)) + \beta\log(1+x) \\ &= -\alpha T x + (\alpha T + \beta)\log(1+x). \end{aligned} \tag{65}$$

By using that $\log(1+x) \leq x - x^2/2 + x^3/3$ for $x \leq 1$, we get

$$\begin{aligned} -\alpha\gamma T + \phi\beta &= -\alpha T x + (\alpha T + \beta)\log(1+x) \\ &\leq -\alpha T x + (\alpha T + \beta)(x - x^2/2 + x^3/3) \\ &= x\left(\beta\left(1 - \frac{x}{2} + \frac{x^2}{3}\right) - \frac{\alpha T x}{2}\left(1 - \frac{2x}{3}\right)\right) \\ &\leq x\left(\beta - \frac{\alpha T x}{3}\right). \end{aligned} \tag{66}$$

The last inequality comes from the fact that $1 - x/2 + x^2/3$ is decreasing in $[0, 1/2]$, and $x = (1-\lambda^2)/2 \leq 1/2$. Choose $T = \frac{3\beta}{\alpha x} + \frac{3C\log(n)}{\alpha x^2}$ for $C \geq 1$, and conclude from (66) that

$$-\alpha\gamma T + \phi\beta \leq -C\log(n).$$

Using the above inequality in (64), we obtain

$$\sum_{s=t+T}^{\infty} \mathbf{P}(\mathcal{A}_{s,\beta}^c \cap \mathcal{E}_{t,\alpha}) \;\leq\; \frac{e^{-\alpha\gamma T+\phi\beta}}{1-e^{-\alpha\gamma T}} \;\leq\; \frac{e^{-C\log n}}{1-e^{-C\log n}} \;\leq\; 2n^{-C}. \tag{67}$$

From (52) and (67), we conclude that

$$\mathbf{P}(\mathcal{E}_{t+T,\beta}^c) \leq 2n^{-C} + \mathbf{P}(\mathcal{E}_{t,\alpha}^c),$$

and this inequality is equivalent to (51). Finally, observe that

$$T \;=\; \frac{3\beta}{\alpha x} + \frac{3C\log n}{\alpha x^2} \;\leq\; \frac{6}{\alpha}\left(\frac{\beta}{1-\lambda} + \frac{2C\log n}{(1-\lambda)^2}\right).$$

□

Using Lemma 5.5, we can prove that in $\mathcal{O}(\log n/(1-\lambda)^2)$ rounds, the number of infected vertices is w.h.p. large enough to apply Corollary 5.4. This is done by applying Lemma 5.5 twice. We first apply it with $t = 0$, $\alpha = 1$ and $\beta = \Theta((\log n)/(1-\lambda))$, which give $T = \Theta((\log n)/(1-\lambda)^2)$, to prove that w.h.p. the infection size does not drop after round T below β before reaching the final target of $\Delta = K \log n/(1-\lambda)^2$. Then we apply Lemma 5.5 again to show that within $T + \mathcal{O}(1/(1-\lambda)^2)$ rounds from the beginning of the process, the infection reaches w.h.p. size at least Δ, which is large enough to apply Corollary 5.4 and obtain the bound (11) of Theorem 2.5.

Lemma 5.6. *Let G be a connected regular n-vertex graph with $1-\lambda \geq \sqrt{2K(\log n)/n}$, where K is the constant form Lemma 5.2. Consider the BIPS process with $|A_0| = 1$ and let T_0 be the first time such that $|A_{T_0}| \geq \Delta = K \log n/(1-\lambda)^2$. Then there exists a large enough constant L such that $T_0 \leq L \log n/(1-\lambda)^2$ with probability at least $1 - 4n^{-3}$.*

Proof. Fix $C = 3$. We first apply Lemma 5.5 with $t = 0$, $\alpha = 1$ and $\beta = 2(\log n)/(1-\lambda)$, which give

$$T = \frac{12(C+1)\log n}{(1-\lambda)^2},$$

and obtain, using the fact that $\mathbf{P}(\mathcal{E}_{0,1}) = 1$, that

$$\mathbf{P}(\mathcal{E}_{T,\beta}) \geq 1 - 2n^{-C}. \tag{68}$$

We apply Lemma 5.5 again, but now β and T, as set above, play the roles of α and t in the lemma. We choose $\beta' = \Delta$, giving $T' = 12(C+K)/(1-\lambda)^2$, and (the values β' and T' play the roles of β and T in the lemma). Then

$$\mathbf{P}(\mathcal{E}_{T+T',\Delta}) \geq \mathbf{P}(\mathcal{E}_{T,\beta}) - 2n^{-C} \geq 1 - 4n^{-C}. \tag{69}$$

The event $\mathcal{E}_{T+T',\Delta}$ implies that there exists a round $s \leq T + T'$ such that $|A_s| \geq \Delta$, so we conclude that with probability at least $1 - 4n^{-3}$, $T_0 \leq T + T' \leq L \log n/(1-\lambda)^2$, for some constant L. □

Lemma 5.6 and Corollary 5.4 imply the bound (11) of Theorem 2.5. It remains to establish that the bound (10) of Theorem 2.5 holds as well. To prove this bound, we first show that w.h.p. the infection has size at least $\Omega((\log n)/(1-\lambda))$ for all rounds $t \geq T$ for some $T = \mathcal{O}((r/(1-\lambda) + r^2)\log n)$. This follows from Lemma 4.6, which we restate in the corollary below for the case of regular graph. Then we repeatedly apply Lemma 5.5 with target sizes increasing geometrically.

Corollary 5.7. *Consider the BIPS process on a connected r-regular graph with n vertices and let $C > 0$ be any positive constant. Then for any $1 \leq k \leq n$ and $t(k) = 4rk + 64(C+2)r^2 \log n$,*

$$\boldsymbol{P}(\exists\, t \geq t(k) :\ |A_t| < k) = \mathcal{O}(n^{-C}). \tag{70}$$

Lemma 5.8. *Let G be a connected regular n-vertex graph with $1-\lambda \geq \sqrt{2K(\log n)/n}$, where K is the constant form Lemma 5.2. Consider the BIPS process with $|A_0| = 1$ and let T_0 be the first time such that $|A_{T_0}| \geq \Delta = K \log n/(1-\lambda)^2$. Then $T_0 = \mathcal{O}((r/(1-\lambda) + r^2)\log n)$ with probability at least $1 - n^{-3}$.*

Proof. Use Corollary 5.7 with $k = \log n/(1-\lambda)$ and $C = 4$ to show that for some $t = \Theta((r/(1-\lambda)+r^2)\log n)$,

$$\mathbf{P}(\mathcal{E}_{t,k}) \ \geq\ \mathbf{P}(\forall s \geq t : |A_t| \geq k) \ =\ 1 - \mathcal{O}(n^{-4}).$$

Now apply repeatedly Lemma 5.5 with $C = 4$ and the parameters (α, β) equal to $(k, 2k)$, $(2k, 4k)$, $\ldots$, $(2^{j-1}k, 2^j k)$, where $j \leq \log_2 n$ is the smallest integer such that $2^j k \geq \Delta$. The value of T in the i-th application of Lemma 5.5 is equal to

$$\frac{6}{2^{i-1}k}\left(\frac{2^i k}{1-\lambda} + \frac{2C\log n}{(1-\lambda)^2}\right) \ \leq\ 6(2C+2)/(1-\lambda) \ \equiv\ \tau,$$

so all applications of Lemma 5.5 give

$$\mathbf{P}(\mathcal{E}_{t+j\tau,\Delta}) \ \geq\ \mathbf{P}(\mathcal{E}_{t,k}) - 2jn^{-4} \ \geq\ 1 - \mathcal{O}(n^{-4}) - 2(\log_2 n)n^{-4} \geq 1 - n^{-3}.$$

Thus with probability at least $1 - n^{-3}$,

$$T_0 \leq t + j\tau = \mathcal{O}((r/(1-\lambda) + r^2)\log n) + \mathcal{O}(\log n/(1-\lambda)) = \mathcal{O}((r/(1-\lambda) + r^2)\log n).$$

□

Lemma 5.8 and Corollary 5.4 give the bound (10) of Theorem 2.5.

Acknowledgments

Research supported by EPSRC grant EP/M005038/1, "Randomized algorithms for computer networks". Nicolás Rivera was also supported by funding from Becas CHILE and by Thomas Sauerwald's ERC Starting Grant 679660 (DYNAMIC MARCH).

References

[1] Noga Alon, Chen Avin, Michal Koucký, Gady Kozma, Zvi Lotker, and Mark R. Tuttle. 2011. Many random walks are faster than one. *Combin. Probab. Comput.* 20, 4 (2011), 481–502. https://doi.org/10.1017/S0963548311000125

[2] Petra Berenbrink, George Giakkoupis, and Peter Kling. 2018. Tight Bounds for Coalescing-Branching Random Walks on Regular Graphs. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*. SIAM, 1715–1733. https://doi.org/10.1137/1.9781611975031.112

[3] Andrei Z. Broder, Anna R. Karlin, Prabhakar Raghavan, and Eli Upfal. 1994. Trading space for time in undirected *s-t* connectivity. *SIAM J. Comput.* 23, 2 (1994), 324–334. https://doi.org/10.1137/S0097539790190144

[4] Flavio Chierichetti, George Giakkoupis, Silvio Lattanzi, and Alessandro Panconesi. 2018. Rumor spreading and conductance. *J. ACM* 65, 4 (2018), Art. 17, 21. https://doi.org/10.1145/3173043

[5] Colin Cooper, Alan Frieze, and Tomasz Radzik. 2009/10. Multiple random walks in random regular graphs. *SIAM J. Discrete Math.* 23, 4 (2009/10), 1738–1761. https://doi.org/10.1137/080729542

[6] Colin Cooper, Tomasz Radzik, and Nicolas Rivera. 2016. The Coalescing-Branching Random Walk on Expanders and the Dual Epidemic Process. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing (PODC '16)*. ACM, New York, NY, USA, 461–467. https://doi.org/10.1145/2933057.2933119

[7] Colin Cooper, Tomasz Radzik, and Nicolás Rivera. 2017. Improved Cover Time Bounds for the Coalescing-Branching Random Walk on Graphs. In *Proceedings of the 29th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '17)*. ACM, New York, NY, USA, 305–312. https://doi.org/10.1145/3087556.3087564

[8] Devdatt P. Dubhashi and Alessandro Panconesi. 2009. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, Cambridge. xvi+196 pages. https://doi.org/10.1017/CBO9780511581274

[9] Chinmoy Dutta, Gopal Pandurangan, Rajmohan Rajaraman, and Scott Roche. 2013. Coalescing-branching Random Walks on Graphs. In *Proceedings of the Twenty-fifth Annual ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '13)*. ACM, New York, NY, USA, 176–185. https://doi.org/10.1145/2486159.2486197

[10] Chinmoy Dutta, Gopal Pandurangan, Rajmohan Rajaraman, and Scott Roche. 2015. Coalescing-Branching Random Walks on Graphs. *ACM Trans. Parallel Comput.* 2, 3, Article 20 (Nov. 2015), 29 pages. https://doi.org/10.1145/2817830

[11] Robert Elsässer and Thomas Sauerwald. 2011. Tight bounds for the cover time of multiple random walks. *Theoret. Comput. Sci.* 412, 24 (2011), 2623–2641. https://doi.org/10.1016/j.tcs.2010.08.010

[12] Uriel Feige, David Peleg, Prabhakar Raghavan, and Eli Upfal. 1990. Randomized broadcast in networks. *Random Structures Algorithms* 1, 4 (1990), 447–460. https://doi.org/10.1002/rsa.3240010406

[13] George Giakkoupis. 2011. Tight bounds for rumor spreading in graphs of a given conductance. In *28th International Symposium on Theoretical Aspects of Computer Science*. LIPIcs. Leibniz Int. Proc. Inform., Vol. 9. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 57–68.

[14] Michael Mitzenmacher, Rajmohan Rajaraman, and Scott Roche. 2016. Better Bounds for Coalescing-Branching Random Walks. In *Proceedings of the 28th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '16)*. ACM, New York, NY, USA, 313–323. https://doi.org/10.1145/2935764.2935791

[15] Alistair Sinclair and Mark Jerrum. 1989. Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. and Comput.* 82, 1 (1989), 93–133. https://doi.org/10.1016/0890-5401(89)90067-9

Appendix

Theorem A.1. ***Chernoff-Hoeffding bound*** *[8, Section 1.6]*
Let $X = \sum_{i=1}^{n} X_i$, where $(X_i)_{i=1}^{n}$ are independently distributed random variables in $[0,1]$. Then, for all $\varepsilon > 0$,

$$\boldsymbol{P}(X > (1+\varepsilon)\mathbf{E}(X)) \leq \exp\left(-\frac{\varepsilon^2}{3}\mathbf{E}(X)\right), \tag{71}$$

and,

$$\boldsymbol{P}(X < (1-\varepsilon)\mathbf{E}(X)) \leq \exp\left(-\frac{\varepsilon^2}{2}\mathbf{E}(X)\right). \tag{72}$$

Theorem A.2. ***Azuma's Inequality*** *[8, Section 5.2]*
Let $(M_k)_{k\geq 0}$ be a discrete-time martingale with respect to a filtration $(\mathcal{F}_k)_{k\geq 0}$. Suppose that for each $k > 0$ there exist a_k, b_k such that

$$a_k \leq M_k - M_{k-1} \leq b_k.$$

Then, for all $n \geq 0$ and $\delta \geq 0$,

$$\boldsymbol{P}(M_n > M_0 + \delta),\ \boldsymbol{P}(M_n < M_0 - \delta) \ \leq\ \exp\left(-\frac{\delta^2}{2\sum_{k=1}^{n}(b_k - a_k)^2}\right). \tag{73}$$